

Lecture 3

01/25/22

• Descriptive Statistics

↳ trying to describe the shape of arbitrary distribution $h(x)$

⇒ based on moments $E(x^n) = \int_{-\infty}^{\infty} x^n h(x) dx$

e.g. mean = $\langle x \rangle = \int_{-\infty}^{\infty} x h(x) dx$

variance = $\langle (x - \mu)^2 \rangle = \int_{-\infty}^{\infty} (x - \mu)^2 h(x) dx$

* NOTE : moments can be biased by outliers

↳ median and σ better judges of location and scale of $h(x)$.

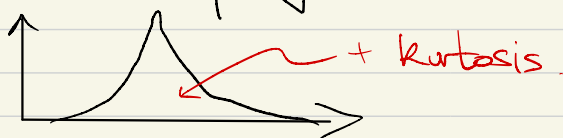
SKWENESS = measures long tail to $+x$

3rd moment



KURTOSIS = measures "peaky-ness"

4th moment



- Sample vs. population statistics

Statistics derived from sampled data are called "sample statistics"

$$\Rightarrow \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

$$s^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2$$

avoids bias, since \bar{x} comes from data also.

NOTE

"s" is not the uncertainty in \bar{x} .

↳ it is an estimator of the scale of $h(x)$

UNCERTAINTIES : $\sigma_{\bar{x}} = \frac{s}{\sqrt{N}}$

$$\sigma_s = \frac{s}{\sqrt{2(N-1)}} = \frac{1}{\sqrt{2}} \sqrt{\frac{N}{N-1}} \cdot \sigma_{\bar{x}}$$

STANDARD ERROR OF QUANTILE : $\sigma_{q_r} = \frac{1}{h_p} \sqrt{\frac{p(1-p)}{N}}$

~~6~~

◦ Univariate Distribution \Rightarrow 1-D.

\Rightarrow whirlwind tour through some important 1D probability distributions.

- * Uniform
- * Gaussian (Normal)

\hookrightarrow confidence levels : $\pm 1\sigma = 68.3\%$
 $\pm 2\sigma = 95.4\%$
 $\pm 3\sigma = 99.7\%$

- * log normal
- * χ^2 distribution
- * χ^2 per dof
- * Poisson distribution
- * Student's t distribution

\Rightarrow The point is that we need to understand how **SIGNIFICANT** our measured statistics are, and how likely it is that they could have been made by chance noise fluctuations.